



US-CMS S&C Drivers 2005

Ian Fisk
CMS Meeting
April 26, 2005



This year we broke the project into 7 execution areas and formed a Project Execution Team (PET)

The PET Tries to meet weekly to align project areas and assess effort issues

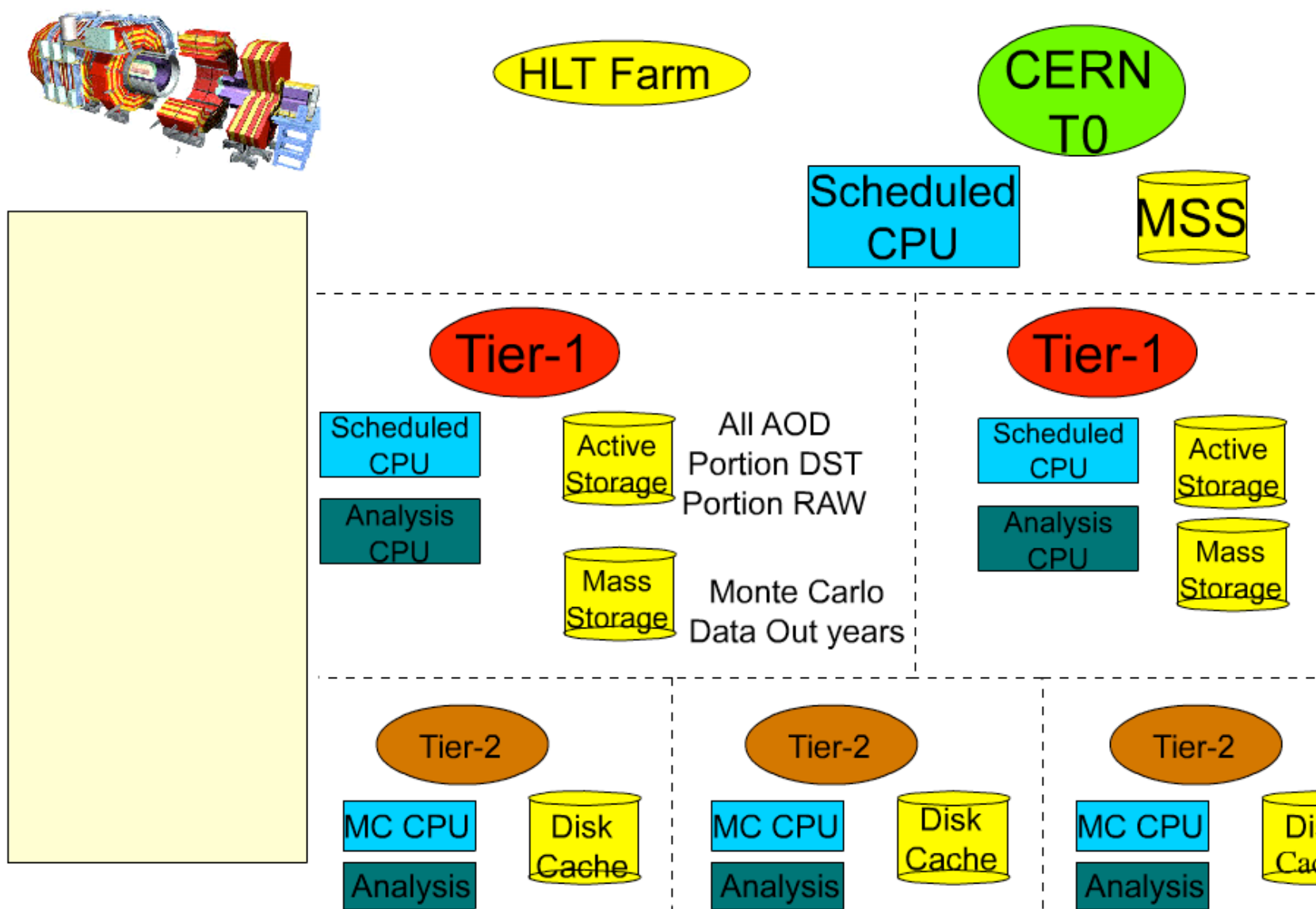
- ➡ WBSI.1 Tier-1 Facility (Led by Jon Bakken)
 - Responsible for development of facilities and facility services
 - Facility Support and Operations
- ➡ WBSI.2 Tier-2 Facilities (Looking for someone)
 - Responsible for Tier-2 Facility Development
 - Tier-2 Integration
 - Tier-2 Operations
- ➡ WBSI.3 Grid Services and Interfaces (Ruth Pordes)
 - Grid Service Development (Accounting services, discovery, info, etc.)
 - Grid Integration and Operations (Formation of Grid3, OSG-x,...)
 - Grid Interoperability



Execution Areas Continued

- ➔ WBS 1.4 Application Services (Currently Lothar, looking for someone)
 - Responsible for data management development
 - Data Placement and Data Bookkeeping
 - Support and operations of current prototypes
- ➔ WBS 1.5 Distributed Computing Tools (Candidate Identified)
 - Responsible for tools for distributed analysis and production
 - Integration with Grid services
 - Operations of the distributed computing environment
- ➔ WBS 1.6 Core Software and Support (Liz Sexton-Kenedy)
 - US Based contribution to core software development (EDM, Arch.)
 - Software packaging, configuration and support
- ➔ WBS 2.0 CCS Liaison (Bob Clare)
 - Working with CCS and tracking US effort contributions to CCS

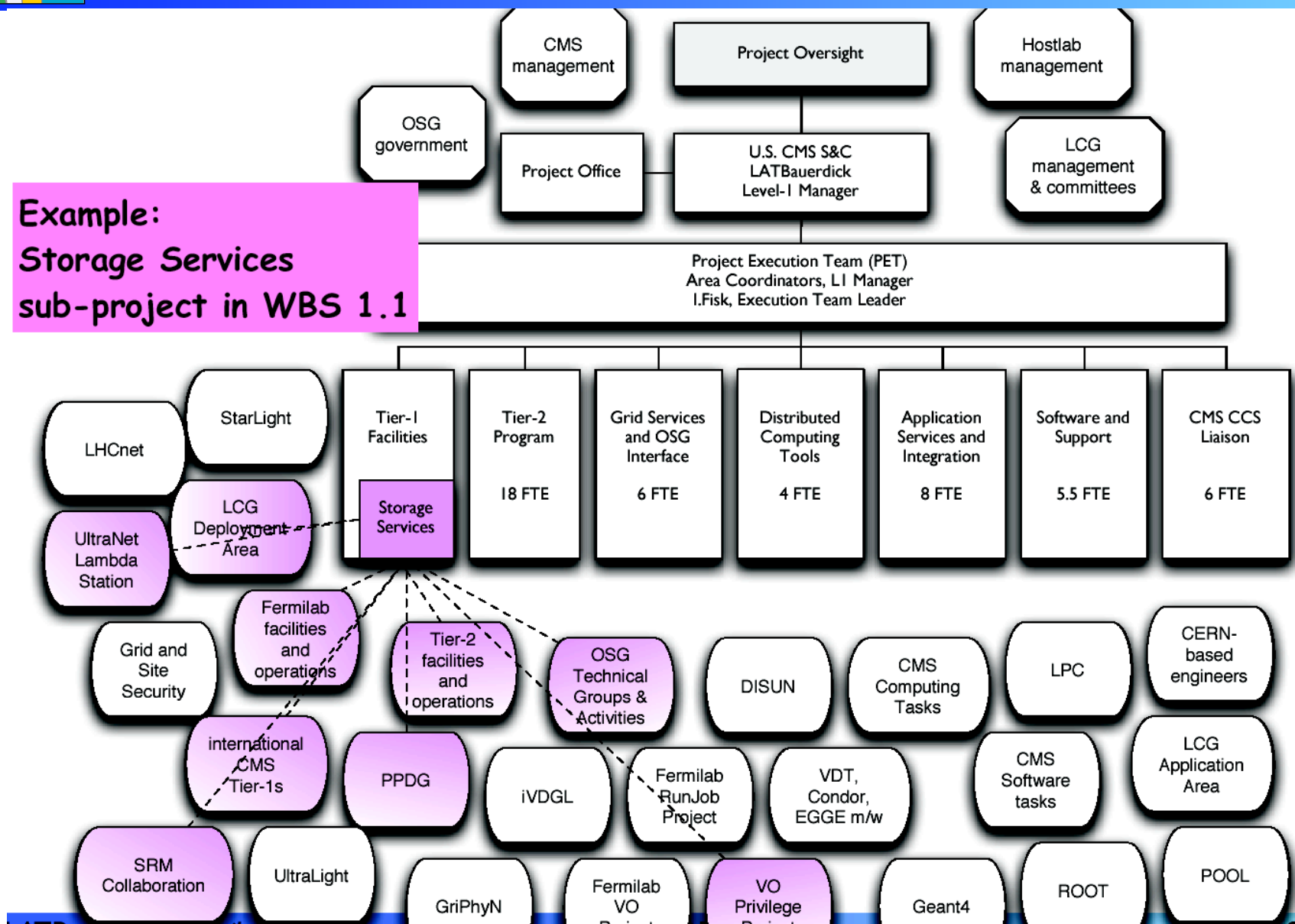
Overview of the Computing System



How Projects Come Together



**Example:
Storage Services
sub-project in WBS 1.1**





Though each Execution Area has an individual WBS with defined scope projects, US-CMS S&C within the scope of International CMS is building a coherent computing system

- ➔ Each year in US-CMS S&C there are a number of high level milestones that we use to track of progress and ensure that we are meeting our obligations
 - Typically aligned with CCS high level milestones, or other external activities important to US-CMS S&C
 - The milestones typically rely on activities from several execution areas for their success
 - Give a higher level view of the project status and schedule
 - Individual WBSs give a lower level view



Driver Milestones (2004 and 2005)



2005 Milestones

➔ Computing TDR

US CMS WBS -- Level 3.4.1.1.3 (/SCPO/Milestones/USCMSSC/USComputingSystems/2005)

http://cmssrv07.fnal.gov:8080/XProject-ifisk/XProject?wbs=3.4.1.1.3&proj=uscms

Lothar's Calendar FNAL Data CDS Agendas - CMS Apple PHP iCalend...sionTracker CMS Page ESNET MeetingPlace CMS Production page UChicago Bookstore

USCMS

THE COMPACT MUON SOLENOID

Home Overview News Links & Resources Contact Search

Software & Computing
Construction / M&O
Physics Analysis

Built Sat Feb 26 23:03:33 2005

WBS	Description
3.4.1.1.3	2005 Milestones
ID: uscms:/SCPO/Milestones/USCMSSC/USComputingSystems/2005 Service Key: current	
3.4.1.1.3.1	2005/05/01: #10368: US-CMS Participation in CMS Computing TDR
3.4.1.1.3.2	2005/07/01: #10370: Data Management Prototype Release
3.4.1.1.3.3	2005/07/02: #10369: Release of Physics TDR Analysis Environment
3.4.1.1.3.4	2005/09/01: #10371: DC06 Production System
3.4.1.1.3.5	2005/10/01: #10365: US-CMS S&C Contribution to Magnet Test
3.4.1.1.3.6	2005/11/01: #10372: Achieve 25% Computing System

Items: 1



US-CMS Computing Milestones for 2005

US-CMS Participation in the Computing TDR (May 2005)

- ➡ CTDR will be released this summer, US-CMS will have several people actively contributing to the document
- ➡ Component and technology evaluations, robust service challenges, and demonstrations feed into the TDR

Release of the Physics TDR Analysis Environment (July 2005)

- ➡ In preparation for final PTDR analysis through the final six months of 2005, US-CMS will finish a round of analysis environment development
 - Improvements in local environment and configuration
 - Development of tools so all US-CMS sites can perform basic analysis steps and accept remote analysis jobs.

Release of Data Management Prototype (July 2005)

- ➡ First release of data management prototype deployed at selected sites
 - Physics queries, data discovery and data cataloging



US-CMS Computing Milestones for 2005

DC06 Production System (September 2005)

- ➡ Release of distributed production components for DC06 simulated events
- Release of an architecture that allows more automated resource discovery, higher level of success job completion, utilization of twice the current number of CPUs.

Magnet Test (October 2005)

- ➡ Participation in framework and event data management evolution

25% Production System (November 2005)

- ➡ Arriving at a system of 25% complexity in 2005
- ➡ Hardware procurements at Tier-1 and Tier-2 centers
- ➡ Configuration management deployment for consistent and centrally supported regional centers
- ➡ Environment development to present a consistent basic environment and services at all regional centers



The Computing Model has a lot of consensus building activities

- ➡ Committees, discussions, documents, presentations, etc.

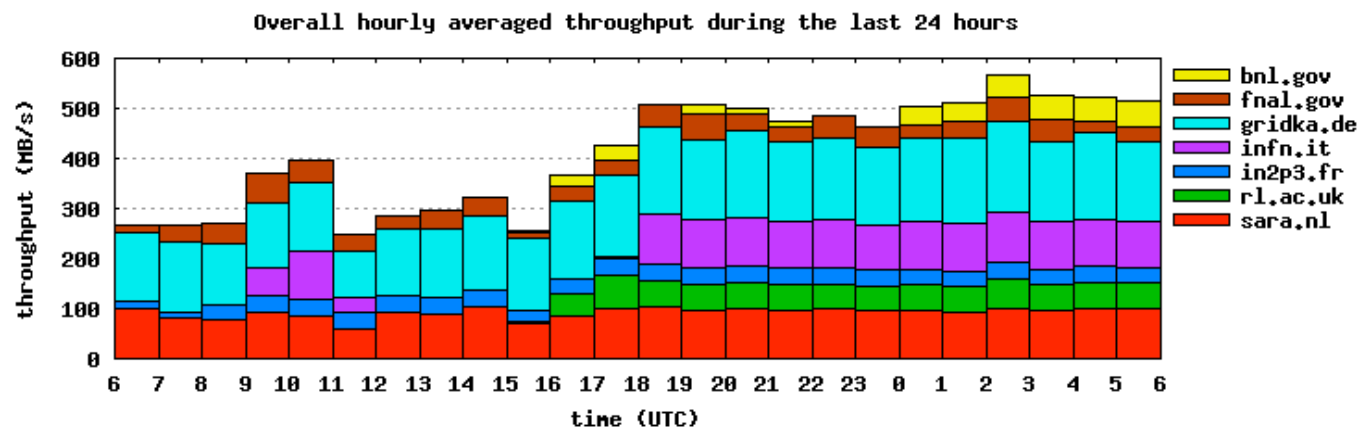
It also has some technical activities that demonstrate feasibility of various aspects of the computing plan

- ➡ LCG Service Challenges: Relies on Activities in
 - WBS 1.1.1.2 Data Storage and Data Access
 - Storage and Interface Development and Support
 - WBS 1.1.1.5 Tier-1 Networking
 - Use of research links and the LHCnet for CERN and Tier-2 Connections
 - LHCNet provides a critical component to test elements at reasonable scale
 - WBS 1.1.2.2 Facility Operations
 - WBS 1.2.3 Tier-2 Operations
 - Expect Participation from Caltech, UCSD, UFL and UW

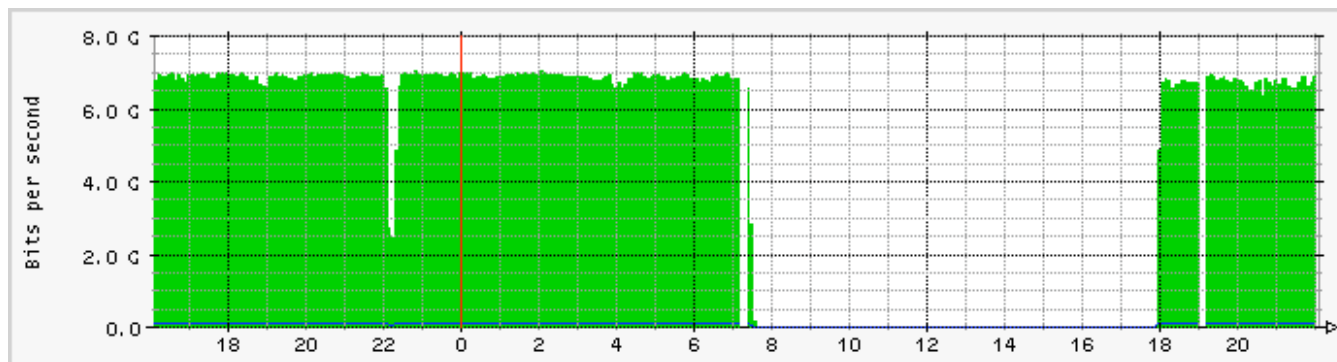


LCG Service Challenges generally progressed well

- ➔ Successfully wrote data to tape from CERN at the agreed upon rate of 50MB/s



- ➔ When constraint to write to tape is removed the rate goes up
 - Transmitted until fiber was cut



Grid Interoperability

- ➔ The resource usage in the CMS Computing model has become more specified over the last 6 months.
- Data is expected to be partitioned at centers, which simplifies match making and simplifies requirements on the grid services
 - Still areas of the CMS Computing model that will rely on high performance grid services that can scale
- ➔ Grid interoperability between LCG and OSG will be required for the efficient interface of the system: VBS 1.3.4

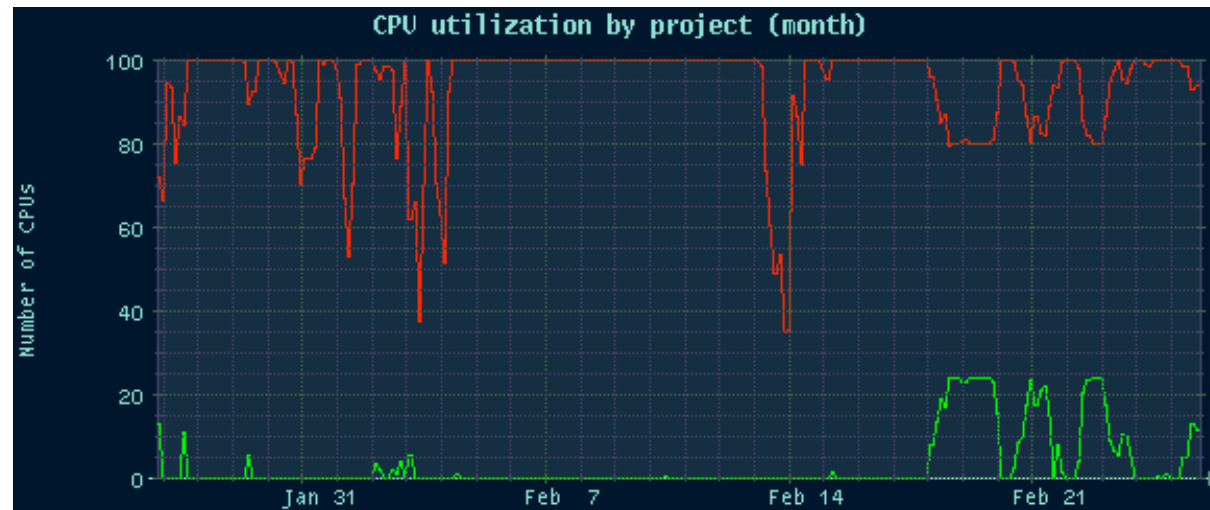
Facility Planning

- ➔ CMS CTDR and CM work tends to update computing requirements at the T1 and T2 centers
- ➔ Budget constraints tend to influence CTDR planning
- ➔ Both requires effort from the T1 program of building and specifying facilities: VBS 1.1.1



Analysis for the Physics TDR has to be finished this year. The User Analysis Facility (UAF) activities so far have been very successful

- ➔ Over 200 registered users, about 20 are active at any given time
- ➔ 56 systems, about 100 batch slots

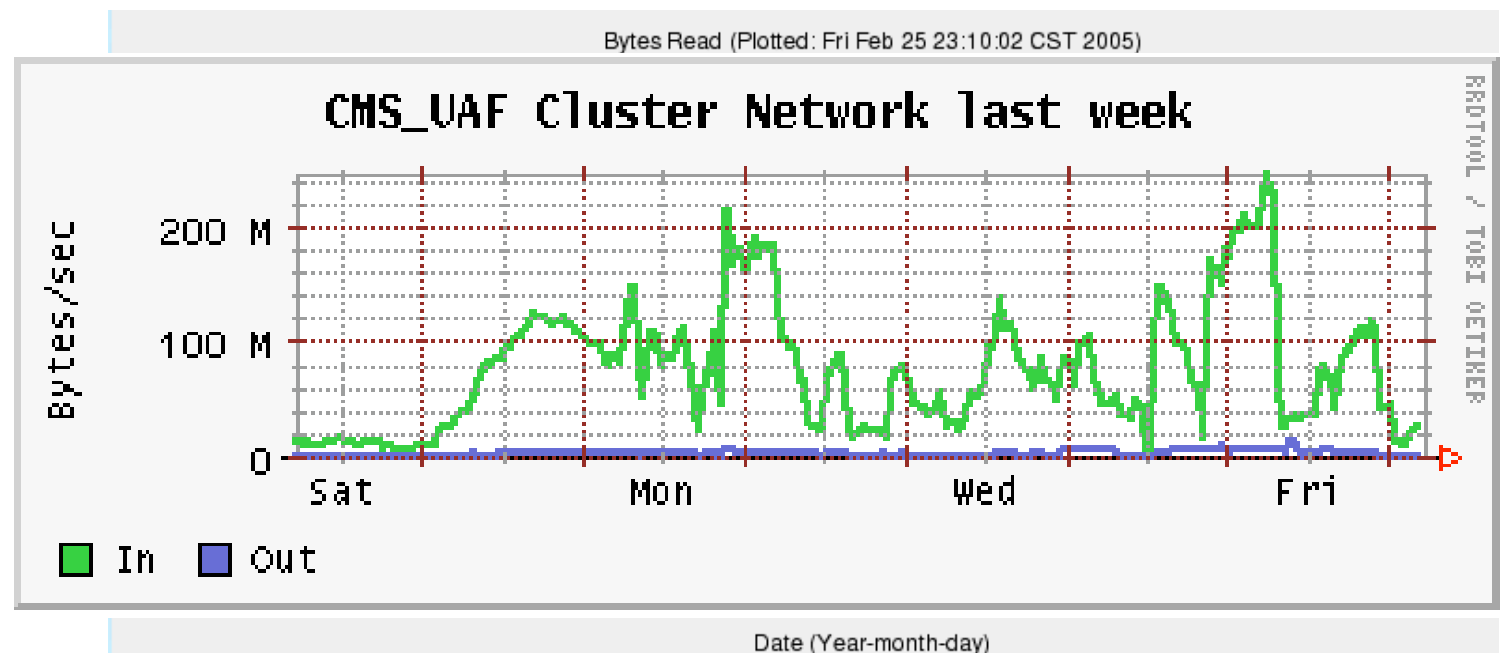


CMS Software is installed, datasets are served, users have scratch space

- ➔ Good examples are published, CMS101 Tutorial has been offered twice, responsive support from the facility staff

The analysis farm has access to ~250 published CMS datasets.

- ➔ Data is published and validated for user community
 - Currently a web page is used, we expect to migrate to Data Management prototype when available
 - <http://www.uscms.org/SoftwareComputing/UserComputing/DataSets.html>
- ➔ Tier-I disk cache demands are increasing





Working with CD



CMS Resources computing and disk resources will double in 2005

- ➡ 300 new worker nodes
- ➡ 64TB of disk chassis
- ➡ To meet the 2006 ramp goals and meet obligations to the user communities we will need to finish the 2005 procurement in a timely manner

Service Challenges will continue

- ➡ SC3 Starts early summer

Operations Continues

- ➡ Need to increase our knowledge base for Grid interfaces, CMS software and services
- ➡ Summer will be a busy time to LPC
 - Data access will ramp up
 - User computing demands will increase



We expect to engage the Tier-2 centers in User Analysis

- ➡ Improvements in data serving and storage
- ➡ Improvements in software environment
 - WBS 1.2.2.2 T2 Service Deployments
 - Effort Estimated at 1 FTE across 4 Tier-2 centers

Software Packaging and Deployment

- ➡ Review and Improve runtime environment at US Sites
 - WBS 1.6.4.3 Core Software and Support
 - Effort estimated at 0.5 FTE for first half of 2005

Improve Remote submission and custom Monte Carlo Services

- ➡ Distributed Computing tools WBS 1.5.1.4
 - DISUN Project is a newly defined program and there is a new leader.

Need to continue to provide support for LPC and UAF users

- ➡ Help desk has been a good resource

Need to formalize our desktop support model for LPC

- ➡ Req for LPC support person is back on the list, but we need to find a good candidate and get someone started
 - Desktop support, application support, environment support, debugging etc.

Need to continue to offer good data access to local analysis users and those that submit through grid interfaces

- ➡ Need to improve integration of facility monitoring and auditing and accounting tools
 - We look forward to working with the accounting project.



Data Management Prototype



DC04 exposed a number of deficiencies in CMS event data management

- ➡ RTAG formed over the summer
 - Report recommended changes in data layout and approach to data storage
 - More stream-like and less database-like

Also prompted the development of a transfer management system for DC04 data distribution

- ➡ System has been used in production since the data challenge and is used to move simulated data between production centers
 - The product was named PhEDEx and has been augmented to improve load balancing and robust transfers
 - Currently being deployed between FNAL and T2s
- ➡ Used to replicate more than 1 TB per day for production data

US-CMS used IFTE of effort in PhEDEx

- ➡ Successful collaboration between US-CMS and UK-CMS



Data Management Prototype (cont.)



The release of the data management prototype is the primary deliverable of the Application Services Execution Area

➡ Dataset Placement service

- Improvements and extensions of the PhEDEx system
 - Wider deployment
 - Interface improvements

➡ Dataset bookkeeping service

- Service to allow query and tracking of datasets for CMS

Initial deployment of prototype at selected facilities by summer

Expect to devote approximately 4FTE of effort in data management development

➡ Project also has a significant international effort

Need to continue to improve the transfer efficiency

- ➡ We get excellent results with the service challenges
- ➡ Just OK results with real production transfers
 - CERN output is a limitation
 - Default CMS data layout appears to be a problem

Need to continue work on making the SRM efficient and reliable

Need to improve the installation and operation of remote dCache + SRM

- ➡ We met with some success at Tier-2 installations, but still dealing with a configuration and operations issues
 - Expect to need to bring up local expertise, but need to improve operations tool also



Data Management and Work Load



Data Management and Workload management are projects of I.4 Application Services and I.5 Distributed Computing Tools

- ➡ Data Management has a prototype scheduled for the summer
- ➡ Workload management working with the Tier-2C program has a prototype in the hands of users expected by the end of the year.
 - Basic prototype expected by the summer

Both represent significant programs of work and need to integrate with

- ➡ Each other
- ➡ Grid Services
- ➡ Facilities

In the CMS Computing Model where the function of some of the resources is largely static and the default location of the dataset is predictable, the Data Management and Workload management are at the heart of the future analysis effort.



Working with CD



Natural progression of our RunJob efforts for some aspects

Some similar philosophies to CAF submission tools

- ➡ Working to maintain some interactive services in grid enabled batch queues.
 - Job monitoring
 - Debugging.



Validation of DC06 Production System

US-CMS has been developing a distributed computing environment for over three years

- ➡ Very large increase in scale and stability
 - 4 sites to more than 20 sites
 - Factor of 10 increase in utilized CPUs for processing

We have leveraged OSG opportunistic computing at universities, national labs, and experiment installations

- ➡ Excellent demonstration of batch computing at the scale achievable

We see that our current production architecture has scaling limitations

- ➡ Pushes jobs down from central submission points
- ➡ Requires significant information exchange for reasonable scheduling decisions
 - Scheduling has manual components
- ➡ Current local job scheduling components are ignorant of VO



Plan of Work for DC06 Production System

We expect to devote approximately 3FTE of effort to achieving a

- ➡ Submission and configuration tools
- ➡ Integration with new OSG Grid services and architectures
- ➡ Evaluation and integration of G-lite Services

Goal is a system capable of making use of a 20% complexity system for simulation generation for DC06

- ➡ Improvements in scale of submission and monitoring and information providing
- ➡ Also should be the basis of DC06 organized CPU usage for reprocessing

We expect to make substantial use of opportunistic resources through the OSG interfaces

- ➡ Storage, Computing, Monitoring, Discovery, Accounting



Magnet Test



January 2006 CMS has a magnet test planned.

- ➡ Slice test of available detector components, DAQ prototypes, and reconstruction software

Magnet Test provides opportunity to work new Event Data Management framework with test data.

- ➡ Provides an integration testbed for on-line side of computing
 - Readout, RAW Data, DAQ, Reconstruction
 - Large scale data challenges like DC04 provide an integration and test environment for the computing environment from T0-T2

The Magnet Test is an end-to-end software test from readout to reconstruction.

The Event Data Management effort in Core Software and Support has 3 total FTE

- ➡ Magnet test is a good way point on a larger program of development



Working with CD



Currently enjoying to support of C++ experts

Project will not last forever, but we would like to formalize our effort abilities while it does last.



Achieving 25% Complexity System



The Tier-1 + Tier-2 centers represent the US-CMS contribution to the experiment computing. 2005 is the first year of the US-CMS hardware ramp in preparation of the start of the experiment. Computing program grow substantially in 06 and 07. In addition to the program funds there are

- ➡ Leveraged additional resources at the universities
- ➡ Opportunistic computing resources provided through the OSG
- ➡ Facility infrastructure resources provided by FNAL and Tier-2 universities

Fine balance between buying late for budgetary reasons and installing sufficient systems to identify and solve scalability problems and develop operational experience.

- ➡ US-CMS plans to ramp to 25% complexity at the T1 and Tier-2 centers in FY05
- ➡ Utilize these resources for pre-challenge event generation and DC06
- ➡ We will devote about 1.5FTE in hardware evaluation and procurement
- ➡ Facility operation and service development is a much larger effort